



Board Game in Physics Classes—a Proposal for a New Method of Student Assessment

Daniel Dziob^{1,2} 

Published online: 27 March 2018

© The Author(s) 2018

Abstract The aim of this study was to examine the impact of assessing students' achievements in a physics course in the form of a group board game. Research was conducted in two groups of 131 high school students in Poland. In each school, the research sample was divided into experimental and control groups. Each group was taught by the same teacher and participated in the same courses and tests before the game. Just after finishing the course on waves and vibrations (school 1) and optics (school 2), experimental groups took part in a group board game to assess their knowledge. One week after the game, the experimental and control groups (not involved in the game) took part in the post-tests. Students from the experimental groups performed better in the game than in the tests given before the game. As well their results in the post-tests were significantly higher statistically than students from the control groups. Simultaneously, student's opinions in the experimental groups about the board game as an assessment method were collected in an open-descriptive form and in a short questionnaire, and analyzed. Results showed that students experienced a positive attitude toward the assessment method, a reduction of test anxiety and an increase in their motivation for learning.

Keywords Assessment methods · Board game · Collaborative testing · Gamification · High school physics

Introduction

In recent years, gamification—which refers to the use of game-based elements, such as game mechanics, esthetics, and game thinking in non-game contexts aimed at engaging people,

✉ Daniel Dziob
daniel.dziob@uj.edu.pl

¹ Smoluchowski Institute of Physics, Jagiellonian University, Łojasiewicza 11, 30-348 Krakow, Poland

² Department of Biophysics, Jagiellonian University Medical College, Łazarza 16, 31-530 Krakow, Poland

motivating action, enhancing learning, and solving problems—has become increasingly popular (Apostol et al. 2013; Deterding et al. 2011). Admittedly, the idea of introducing games in teaching is not a new concept. People have been using digital games for learning in formal environments since the 1960s (Ifenthaler et al. 2012; Moncada and Moncada 2014). However, the term of gamification was coined only a few years ago, and since then has been gaining more and more popularity (Dicheva et al. 2015; Sung and Hwang 2013). The benefits of gamification (or, in more broad terms, game-based learning, e.g., Ifenthaler et al. 2012) in educational contexts are widely described in the literature. Among them are increasing student intrinsic motivation and self-efficacy (Banfield and Wilkerson 2014; Seaborn and Fels 2015), motivation effect and improvement of the learning process (Dicheva et al. 2015; Sadler et al. 2013), as well as improving the positive aspects of competition (Burguillo 2010; Conklin 2006).

Games can reinforce knowledge and bridge the gap between what is learned by creating dynamic, fun, and exciting learning environments (Royse and Newton 2007). They are a powerful teaching strategy, and they challenge and motivate students to become more responsible for their own learning (Akl et al. 2013). However, this requires having the game to be well-designed and structured clearly with a framework that provides effective outcomes (Allery 2004). The review presented in Dicheva et al. (2015) suggests that early adopters of gamification are mostly Computer Science/IT educators. This is in line with the rising popularity of computer games, which have become prominent in the last decade. Nowadays, many articles can be found, in which using computer games in the teaching process are introduced and evaluated (Eskelinen 2001; Ko 2002; Rieber and Noah 2008). Nevertheless, not all of them are proper for school circumstances. Zagal et al. (2006) points out that some of the designed games are highly opaque and complex in rules, and did not include players collaborating to play the game: Therefore, these games did not affect students peer-learning. Through peer collaboration, students build on each other's knowledge to develop new attitudes, cognitive skills and psychomotor skills (Adams 2006; Damon and Phelps 1989). The same authors suggest that for such a purpose board games could be used due to their transparency regarding the core mechanics. Moreover, board games provide the teachers with an opportunity to guide or direct children to meet specific educational goals by extending their learning during and after playing the game (Durden and Dangel 2008; Wasik 2008). Teachers can also facilitate communication amongst children, build understanding about games, discuss concepts, and provide feedback to one another (Griffin 2004).

Board games are also successfully used in early childhood education (Ramani and Siegler 2008; Shanklin and Ehlen 2007) as a pedagogical tool that reinforces a positive environment for learning (Dienes 1963). Games also appear to build positive attitudes (Bragg 2003) and self-esteem, and enhance motivation (Ernest 1986). They have been found to be also effective in promoting mathematical learning (Bright et al. 1983), mathematical discussion (Ernest 1986; Oldfield 1991), social interaction (Bragg 2006), and risk taking abilities (Sullivan 1993). Some types of board games were also used in medical education and have been found as useful methods for conveying information and promoting active learning (Neame and Powis 1981; Richardson and Birge 1995; Saunders and Wallis 1981; Steinman and Blastos 2002). In the present study, a board game—a competitive game between groups of students in a classroom—was used as an assessment in order to examine if it could increase high school students' achievements and retention of knowledge in physics.

To assess student achievements in general, and as a result of the board game specifically, there are two main formats of assessment distinguished and widely discussed in the literature,

namely formative and summative assessment (Harlen and James 1997; Wiliam and Black 1996). In general, formative assessment is carried out throughout a unit (course, project) and its purpose is to provide feedback to students about the learning process. Summative assessment is given at the end of a unit (course, project) and is used to summarize students' achievements usually in the form of grades (Harlen and James 1997; Looney 2011; McTighe and O'Connor 2005). Even though summative assessment could be performed in many ways, some authors pointed to a lack of post examination feedback for students as a weakness (Leight et al. 2012; Talanquer et al. 2015). In our study, the board game was used essentially as a tool for summative assessment, although it also includes some elements of formative evaluation. Such a combination was dubbed by Wininger (2005) as a *formative summative assessment*. It entails reviewing exams with students so that they get feedback about their knowledge comprehension. One example of this approach is collaborative testing that aims to give students an opportunity to work in groups during or at the end of an exam (Guest and Murphy 2000; Lusk and Conklin 2003). Research has shown that there are many benefits to collaborative testing. These are described in detail by Duane and Satre (2014), Gilley and Clarkston (2014), Kapitanoff (2009), and based also on literature about the positive impact of group testing (Millis and Cottell 1998; Michaelson et al. 2002; Hodges 2004) and peer-learning (Slusser and Erickson 2006; Meseke et al. 2008; Ligeikis-Clayton 1996), which are parts of collaborative learning. Among others, the most important benefits of collaborative learning are increasing students' achievements (Bloom 2009; Haberyan and Barnett 2010), reduction of test anxiety (Zimbardo et al. 2003), improvement of critical thinking ability (Shindler 2003), and collaboration skills (Lusk and Conklin 2003; Sandahl 2010).

The assessment in the form of a game employed in the current research is based on the authors' previous experiences and research (Dziob et al. 2018). It evaluates not only the content matter knowledge itself, as in typical tests, but combines a few different aspects together, as schematically shown in Fig. 1. It assesses the relationship between content knowledge and everyday life, as well as socio-historical context. Moreover, it gives the opportunity to assess research skills required to conduct experiments. The form of the board game enables development of social and entrepreneurial skills in the form of a *challenge-yourself* competition, which allows students to surpass individual limitation (Doolittle 1997).

This study reports on the efficacy of assessing students' knowledge by means of a group board game approach and measuring its effects on students' learning outcomes. The research questions are as follows:

- 1) What is the effect and influence of the board game assessment on student learning outcomes when compared with student prior results in physics?
- 2) What is the effect and influence of the board game assessment on student learning outcomes when compared with a traditional teaching approach?

Methodology

Participants

The research was conducted on a group of 131 students in total from two high schools in Poland. Students were divided into experimental (of $n = 37$ and $n = 36$ in school 1 and 2, respectively) and control groups ($n = 31$ and $n = 26$). Each group was taught by the same

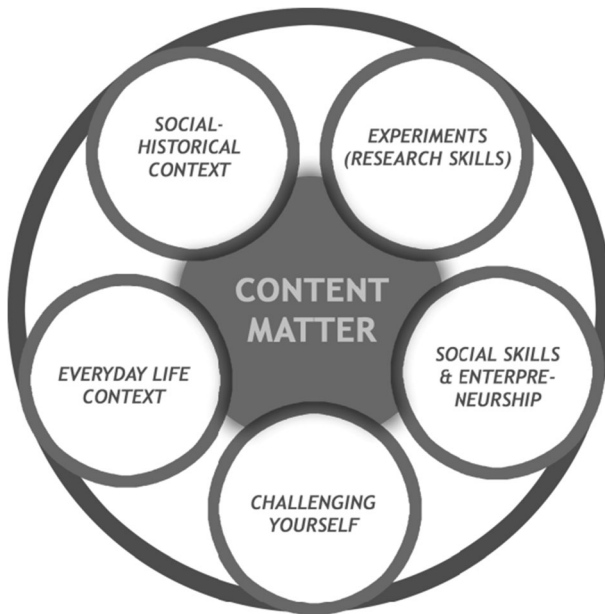


Fig. 1 Assessment strategy components. In addition to the content matter knowledge, all other expressed elements were involved in the assessment process. Own work

teacher and followed the same curriculum during their education. Just before the experiment, the students had accomplished a 25-h course on vibrations and waves (in school 1) and on optics (school 2). After finishing the unit, experimental groups took part in the assessment in the form of a group board game (described below, hereinafter the intervention) and 1 week after in a traditional test. Students from control groups participated only in the traditional test the same as experimental group, but without the intervention. In each group, the ratio of males to females was similar (3:2).

Intervention

The section below contains a detailed description of the intervention: a game which students from experimental groups played once at the end of the unit together with the evaluation process. The description includes procedure and examples of questions used in students' assessment.

Intervention Organization

The game lasted approximately 2.5 lesson hours (approx. 110 min). At the beginning, students were divided randomly into groups of 4 to 5 people each, and asked to take seats around the game board table. Each group began with questions concerning some physics phenomena and students moved their tokens (one per group) forward by the number of right answers or correctly named concepts. At the end of the game (when allocated time ended), students were asked to fill in a self- and peer-assessment questionnaire. At each stage of the game after students had attempted, a scientifically accepted answer to each question was provided together with a proper explanation by students or, if needed (when students didn't pass), by

the teacher. Thus, this approach allows the teacher to immediately rectify and clarify students' misconceptions.

Game Board—Organization

The game board consisted of a circular path, and the participants moved their group token along this path. The path was made up of a random assortment of five potential categories, or activities, to land on. These included physics phenomena charades, famous people, short answer questions, multiple-choice questions, and simple experiments. All questions required the students to perform different types of activities and allow them to obtain a different number of points. Because the number of points obtained at each stage was identical with number of spots the token was moved, the scoring system was identical with the movement system as in typical board game. Additionally, there were also two special lines on the board. Whenever any group reached one of them, the members of both groups received special algebraic tasks or complex experimental tasks to solve. Figure 2 presents the game board with the fields of different type indicated on it.

Physics Phenomena Charades

Upon reaching this field, one representative of a given group received six cards with names of various physics phenomena related to waves and vibrations (school 1) or optics (school 2; see Fig. 3). Their aim was to describe each concept, without using the words given, so that the rest of the team could guess the name. The time for this task was limited to 1 min (measured by a small

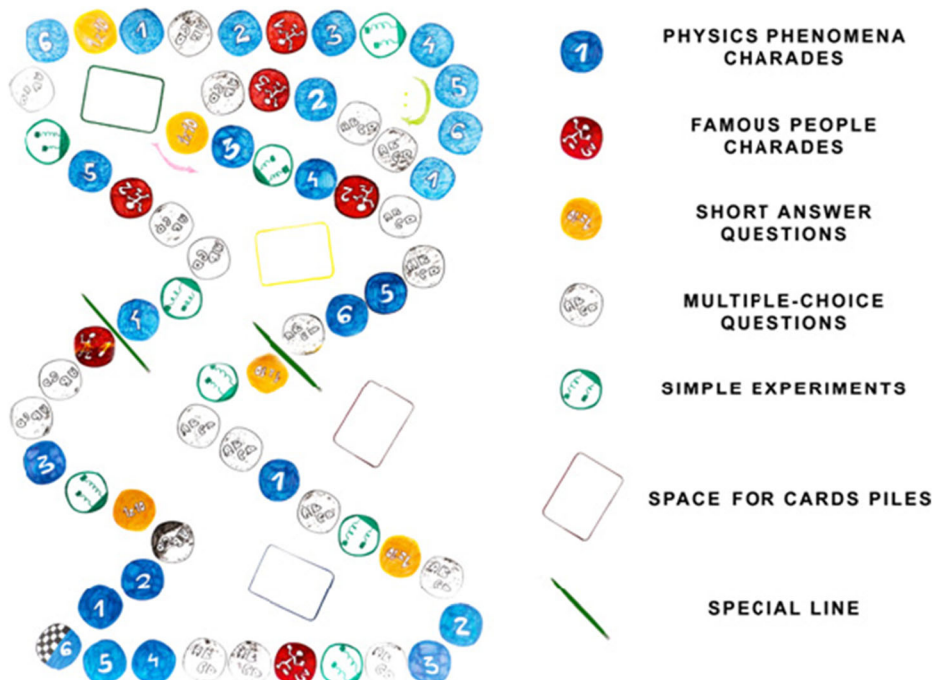


Fig. 2 The design of the board game

hourglass). After the end of the round, tokens were moved forward by the number of fields equal to the number of correctly guessed charades.

Famous People Charades

These questions were similar to the previous ones, but they related to important people connected with the concepts of waves, vibrations, and acoustics (physicists, musicians, etc.) or optics (Fig. 4). The scoring system was identical to the one employed in the physics phenomena charades.

Short Answer Questions

The short answer questions differed with respect to their level of difficulty, but usually they required only a true/false answer (see Fig. 5). The questions were asked by the teacher and the time of each group's round was 1 min. Within that time, all members of the currently active group could answer as many questions in a row as they managed, without passing their turn to another group. If the provided answer was wrong, the next group took over and had an opportunity to answer other questions. At the end of each round, the groups moved their token forward by the number of the correctly answered questions divided by 2 and rounded up.

Multiple-Choice Questions

Upon reaching a field of this category, a group received multiple-choice questions related to scientific reasoning (Fig. 6). Students had to point out the correct answer and provide comprehensive argumentation for their choice. By providing the right answer together with the correct explanation, students could move forward by 2 fields on the board. Otherwise, no move was allowed.

Simple Experiments

Upon reaching a field of this category, the students had to conduct some simple experiments in order to prove relevant phenomena (Fig. 7). The equipment necessary for each experiment, with some extra materials among them, were available to students. An important part of the task was the necessity to make a decision which objects were essential. The other groups taking part in the game were enabled to ask the currently active team detailed questions about the conducted experiment and ask for additional explanations. Having carried out the experiment and addressed the questions properly, the group was allowed to move forward by 2 fields.

Waves and vibrations:	Optics:
<i>Amplitude</i>	<i>Incident ray</i>
<i>Isochronism</i>	<i>Diffuse reflection</i>
<i>Wave front</i>	<i>Interference pattern</i>
<i>Superposition principle</i>	<i>Ray diagram</i>
<i>Fundamental frequency</i>	<i>Focal point</i>
<i>Ear canal</i>	<i>Light absorption</i>

Fig. 3 Examples of physics phenomena charades

Waves and vibrations:	Optics:
<i>Ernst Chladni</i>	<i>Willebrord Snell</i>
<i>Robert Hooke</i>	<i>Zacharias Janssen</i>
<i>Christian Huygens</i>	<i>Thomas Young</i>
<i>Alexander Graham Bell</i>	<i>Hans Lippershey</i>
<i>Wolfgang Amadeus Mozart</i>	<i>Galileo Galilei</i>

Fig. 4 Examples of famous people charades from games on waves and vibrations and optics

Algebra Tasks

When one of the groups reached the first special line on the game board after the end of the round, all competing groups simultaneously received three algebra tasks. They had 10 min to solve them. For accomplishing this task, each group could receive a maximum of 4 points and moved their token forward by 4 fields. Incorrect or incomplete solutions, assessed by the teacher, reduce the amount of points.

Experimental Task

When one of the groups reached the second special line on the board at the end of the round, all competing groups simultaneously received one experimental task, which was neither discussed nor solved during any previous class. The students had to come up with an experimental design to examine the effect of damping the movement of the harmonic oscillator (school 1) or examine the surrounding medium refractive index on the glass lens focal length (school 2). The groups received special worksheets prepared in accordance with an inquiry-based methodology. Students had to formulate a proper hypothesis, describe the plan of the experiment, draw the experimental setup, write down their observations, analyze the results, and draw the conclusions. This part took up to 20 minutes. For this task, students could receive a maximum of 10 points.

Instruments and Data Collection

Former Achievements

Before the intervention, students from each group were tested individually in four tests throughout the school year; on kinematics, energy, gravitation, and rigid body rotational motion (school 1);

Waves and vibrations:
1. <i>In which state of matter does a sound wave have a maximum speed?</i>
2. <i>Mechanical waves can be divided into transverse and ...</i>
3. <i>Can harmonic motion be described by a tangent function?</i>
4. <i>A standing wave does not change its frequency over time. True or false?</i>
Optics:
1. <i>If Ann is standing 3 feet in front of a plane mirror, how far from her will her image be located?</i>
2. <i>The white light consists only of red, green and blue colors. True or false?</i>
3. <i>Can a real image be obtained in the concave mirror?</i>
4. <i>Lenses can be divided into converging and...</i>

Fig. 5 Examples of short answer questions

Waves and vibrations:

Q: Which oscillatory motion is not a harmonic motion?

- A. motion of reflecting ball*
- B. motion of the weight on the spring*
- C. vibrating motion of the liquid in the tube-shaped U*
- D. movement of the tube floating on the liquid surface after partially submerged and released*

Why?

Optics:

Q: Which quantity always changes when light is passes from one medium to another with a different speed of light within?

- A. Frequency*
- B. Color*
- C. Direction*
- D. Wavelength*

Why?

Fig. 6 Examples of short answer questions

and electrostatics, current, magnetic field, and induction (school 2). They comprised mixed problems: content knowledge and scientific reasoning tasks, multiple-choice, open-response, and algebra problems. Tests were the same for experimental and control groups. The average of each student's percentage results on the four tests was used to measure his/her achievement prior to the game, henceforth called *average former achievements* and denoted as *FA*.

Assessment Questionnaires

When the game ended, each student was asked to fill in individually two questionnaires of self- and peer-assessment in order to evaluate themselves and other fellow players from the same group under various sides. Each questionnaire was composed of eight questions designed on a 6-point Likert scale. Half of the questions focused on the students' communication skills, while the rest on subject matter contribution. In Table 1, the self-assessment questionnaire is presented. The peer-assessment questions were designed in the similar way.

Evaluation Process

The questionnaire-based assessment results were included in the final score according to the author's own approach presented below and described in detail in Dziob et al. (2018).

Waves and vibrations:

- 1. Create a standing wave in the air.*
- 2. Experimentally demonstrate the superposition principle.*
- 3. Demonstrate that a wave does not cause matter to move towards it distribution.*

Optics:

- 1. Experimentally demonstrate the law of reflection.*
- 2. Show the refraction of light.*
- 3. Demonstrate that the light always travels via a straight line.*

Fig. 7 Examples of simple experiments tasks

Table 1 Student self-assessment questionnaire

Question	1–6 scale
Were you involved in the group work?	Communication skills
Did you communicate adequately with the group?	
Did you take part in the discussion on the problem?	
Did you take into account the opinion of other students?	Subject matter contribution
Did you prepare for the test beforehand?	
Did you participate in solving problems and tasks?	
Did you have sufficient knowledge to solve the issues?	
Did you contribute to the final result of the group?	

1. The mean average score was calculated based on the “subject matter contribution” and, separately, “communication skills” points in the self-assessment results (S).
2. The average score was calculated based on the “subject matter contribution” and, separately, the “communication skills” points ascribed to the student by the other members of the group (the peer-assessment, P).
3. At the end, the “subject matter contribution” and “communication skills” scores were obtained separately as follows:
 1. If $|S - P| \leq 1$ (a consistent evaluation): take P as the final score
 2. if not (an inconsistent evaluation): take $P - 0.5$ as the final score

The percentage score for each team was calculated by dividing the number of points (number of fields) accumulated by the group by the maximum number of points available to obtain. The final overall score given to each student consisted of three parts:

1. the group common percentage result from the board game—with the weight of 0.5,
2. the questionnaire-based assessment percentage result for the “subject matter contribution”—with the weight of 0.3,
3. the questionnaire-based assessment percentage result for the “communication skills”—with the weight of 0.2.

The final score for each student after the game, calculated according to the algorithm above and expressed in a percentage form, is henceforth referred to as *game score* (GS).

Post-test

An unannounced post-test was conducted in the experimental groups 1 week after the game. The same test was given to students from the control groups, just after finishing the unit. It was prepared in a traditional written form. There was neither a review of the relevant content knowledge during regular classes nor a post-game discussion of the game problems and results before this test. The *post-test* (PT) score is expressed in percentage terms.

Students’ Opinions Questionnaire

Students from the experimental groups received an anonymous short evaluation questionnaire a week after the game (just after the post-test). It consisted of six questions asking “How the knowledge assessment method influences your...”, and each answered on a linear point scale

ranging from -5 to $+5$, with the numbers indicating the most negative (-5), through none (0) to the most positive ($+5$) impact. The evaluated aspects covered pre-test preparation, engagement in team work, answer difficulty, test anxiety, final acquisition of knowledge, and motivation for future learning. It was also a space for students to present opinions on the game. Exact questions are presented in the results section together with students' answers.

Data Analysis

Basic Statistical Analysis

Below, a statistical analysis of the data is carried out, firstly for the experimental groups, and then in comparison with the control groups. In Table 2, we present basic descriptive statistics and empirical distributions (in the form of histograms, with normality tested by the Shapiro-Wilk tests) for each set of results, i.e., the FA, GS, and PT, for both experimental groups. The numbers 1 and 2 in superscripts indicate the schools.

All examined variables have normal distribution. Student t test showed that the differences among each variable means are statistically significant (in each case p value < 0.05). This allowed for comparison of the students' results in different tests. On average, the students from experimental groups scored 47%/59% (school 1/school 2) in the former test, 70%/80% in the game, and 58%/68% in post-test. An increase of almost 23 percentage points (pp.) between FA and GS in both experimental groups might emerge as the result of student cooperation during the game. The PT results are lower than GS. However, they are still statistically significantly higher than the FA results ($p < 0.05$), which may suggest a positive impact of game-based assessment on students' achievements. It should be noticed, however, that at each stage the results of students from the first school are lower than those from the second. This is consistent with the author's observation about the educational standards in each school. Therefore, in what follows, both groups will be analyzed independently, in comparison to adequate control groups from the same schools.

In both schools, control groups were formed from the students who studied with the same teacher and who completed the same courses. In Tables 3 and 4, basic descriptive statistics for control and experimental groups former achievements (FA) and results in post-test (PT) are presented. In each school, average former achievements of the students from both the

Table 2 Basic statistics of the results obtained by students from experimental groups

Variable	Characteristics						
	Mean	95% Confidence interval for mean	Median	Lower quartile	Upper quartile	Standard deviation	
FA ¹ [%]	47.05	(44.38; 49.72)	45.41	40.31	54.67	8.00	School 1
GS ¹ [%]	69.74	(66.46; 73.03)	69.71	64.58	77.47	9.86	
PT ¹ [%]	58.44	(54.29; 62.56)	57.58	48.48	69.70	12.44	
FA ² [%]	58.58	(53.07; 64.07)	55.37	46.68	68.54	16.25	School 2
GS ² [%]	80.01	(76.27; 83.76)	79.81	71.56	87.84	11.07	
PT ² [%]	67.76	(62.56; 72.97)	68.36	56.78	76.82	15.38	

Superscripts indicate the schools

FA average former achievements, GS the final score in the game, PT the result in the post-game test

experimental and the control group are similar. As proved by the t test (each data are normally distributed), there are no statistically significant differences ($p > 0.05$) between FA in experimental and control group within one school.

The former achievements and post-test results within the control groups were tested in the same way. Results indicate ($p > 0.05$) that there is no statistically significant difference between former achievements (FA^C) and post-test results (PT^C) in control groups. It implies that the post-test can be considered as a reliable tool, neither harder nor simpler than the former test. It allows the comparison of the PT results between experimental and control group. In school 1, the difference between mean results is close to 8 pp. ($p = 0.0000$), and in school 2, it is a little bit above 10 pp. ($p = 0.0003$). It clearly shows that experimental groups obtained statistically higher results in PT than their colleagues from the control groups. In other words, students from the experimental groups gained significantly more knowledge than their colleagues from the control groups.

Students' Opinions

Students' opinions about the board game were collected just after the post-test, but before providing them with the information about their final marks. Participants filled in a questionnaire (on -5 to $+5$ scale, where 0 means no impact) and expressed their comments anonymously in an open, descriptive form. Mean results for the questionnaire questions in both schools are provided in Table 5.

Because the answers for each were normally distributed (tested by Shapiro-Wilk test), the comparisons of the H_0 : mean against zero was calculated using the Student t test. The test showed ($p < 0.05$) that in each question students' answers differ significantly (were higher or lower) than "0" value, which means no impact. In other words, for each question, students state significant influence of the board game on the tested issue. Students in both schools judged the assessment in the form of a group board game beneficial for their preparation, and pre-test preparation was marked by students from each school as positive. This means that students would spend more time preparing for the game, as opposed to spending time on preparation for a traditional test. Both experimental groups agreed that the level of engagement of their team-mates was high and that answering questions was easier than in traditional, individually taken tests. It corresponds with the students' opinions that this new form of assessment prompts them to give answers, even if they feel uncertain about their correctness. Anxiety during the test was assessed at -3.1 and -2.3 in both experimental groups, respectively, which means that this form of assessment reduces the

Table 3 Basic statistics of the students from the first school

Variable	Characteristics						
	Mean	95% Confidence interval for mean	Median	Lower quartile	Upper quartile	Standard deviation	t test
FA^1 [%]	47.05	(44.38; 49.72)	45.41	40.31	54.67	8.00	$p = 0.3120$
FA^{1C} [%]	49.43	(45.31; 53.55)	48.28	42.07	58.12	11.23	
PT^1 [%]	58.44	(54.29; 62.56)	57.58	48.48	69.70	12.44	$p = 0.0000$
PT^{1C} [%]	50.67	(46.43; 54.92)	49.34	41.03	61.03	11.57	

Superscripts differentiate experimental (1) and control group (1C) within first school

FA average former achievements, PT the result in the post-game test

Table 4 Basic statistics of the students from the second school

Variable	Characteristics						
	Mean	95% Confidence interval for mean	Median	Lower quartile	Upper quartile	Standard deviation	<i>t</i> test
FA ² [%]	58.58	(53.07; 64.07)	55.37	46.68	68.54	16.25	$p = 0.6539$
FA ^{2C} [%]	56.81	(51.21; 62.41)	56.49	51.78	64.11	14.16	
PT ² [%]	67.76	(62.56; 72.97)	68.36	56.78	76.82	15.38	$p = 0.0003$
PT ^{2C} [%]	57.65	(51.43; 63.87)	58.52	44.04	65.76	15.72	

Superscripts differentiate experimental (²) and control group (^{2C}) within second school

FA average former achievements, PT the result in the post-game test

anxiety normally associated with traditional exams. Students also indicated, both in the questionnaires and open opinions, that the board game improved their final level of knowledge. Students also felt motivated by the game to continue learning.

A few examples of the opinions are presented below:

- Student A:

This is a good option to test for people who are weaker in calculation. Not everyone is able to solve a complex task, but anyone can learn theory.

- Student E:

This form of the test was very good, because you could learn also during the test. It teaches cooperation in the way you could have fun.

- Student K:

I think that we learned and invented more during this game than during a written test. It was a very good possibility for integration.

- Student O:

Each group should get all kinds of questions. Then it would be more fair. Questions should be more focused on physics, without connections to history.

Table 5 Mean results for each question in the questionnaire for both experimental groups

How the form of the assessment influences your		School 1	School 2
1.	Pre-test preparation	3.2 (1.1)	2.8 (1.6)
2.	Engagement into team work	3.6 (0.9)	3.7 (0.8)
3.	Easy of answering	2.9 (1.8)	3.4 (1.2)
4.	Test anxiety	−3.1 (1.6)	−2.3 (2.1)
5.	Final acquire of knowledge	2.4 (1.8)	3.9 (1.2)
6.	Motivation for future learning	3.5 (0.9)	3.4 (1.7)

Numbers in brackets denote standard deviations

The vast majority of students' opinions were positive and enthusiastic. A few of them used the feedback to provide helpful and insightful comments for improving the assessment. In the discussion section, we relate them to the findings commonly presented in the literature on collaborative testing and gamification.

Discussion

The main purpose of this research was to investigate the influence of assessing students' achievements in the form of a group board game in comparison to their former achievements and traditional tests. The first important finding is a statistically significant increase in students' achievements in the game in comparison to their former achievements. This result is consistent with research on the positive impact of collaborative testing, which shows that students' results obtained in collaborative taken exams are higher than in individual ones (Bloom 2009; Haberyan and Barnett 2010; Kapitanoff 2009; Lusk and Conklin 2003). Some authors controvert, however, the ability of collaborative testing to improve content retention (Leight et al. 2012; Woody et al. 2008), pointing out that only their performance during the collaborative exam is higher. Our second results addressed this problem. We found that students from experimental groups gained higher results in the post-test taken 1 week after the game with respect to the results obtained by the control groups. In other words, the students assessed by the game obtained not only high performance in the game but also in a knowledge test taken after the game. This finding is encouraging with respect to other research that shows improvement in students' achievement after the collaborative exam in the long run (Cortright et al. 2003; Jensen et al. 2002; Simpkin 2005). The results show also that the assessment method is efficient independently of the level of students' performance.

The students' opinions were encouraging and supported findings in the literature. Board games can be perceived as a form of activity in which group work skills are exploited and play an essential role in accomplishing tasks (Dallmer 2004; Kapitanoff 2009; Lusk and Conklin 2003; Sandahl 2010; Seaborn and Fels 2015; Shindler 2003). Some researchers (Dicheva et al. 2015; Sadler et al. 2013) suggested that *gamification* could improve the learning process, which can be inferred from the increase in students' results in post-test. By playing the game, the students learn to listen to everybody else's answers, provide fellow players with their know-how, and respond to ideas proposed in discussions. According to Hanus and Fox (2015) and Jolliffe (2007), the above can stimulate knowledge assimilation. In the students' opinions expressed in the questionnaire and open-descriptive form, the board game assessment has a positive impact on their motivation and social interactions, which also corresponds to the literature findings (Banfield and Wilkerson 2014; Bragg 2006; Seaborn and Fels 2015). Furthermore, the assessment in the form of a game induces far less test anxiety by giving students a sense of being supported by the other team members (Banfield and Wilkerson 2014; Kapitanoff 2009; Lusk and Conklin 2003; Sandahl 2010; Zimbardo et al. 2003). Similar results can be found in other research, in which results from student's attitude surveys confirm that collaborative testers have more positive attitudes towards the testing process in general compared to students who take assessments individually (Bovee and Gran 2005; Giuliadori et al. 2008; Meseke et al. 2009). Finally, an active involvement in the self- and peer-assessment process may increase the students' self-assurance and adequate self-esteem (Hendrix 1996), thereby enhancing retention of knowledge (Sawtelle et al. 2012).

Comments on Organization of the Game

Preparation of the board game has a few important aspects, which should be described in order to give the reader the impression of how to adapt the idea to her/his own purpose. The most important is to decide on the topic, which should give considerable benefits for assessing knowledge in a non-standard form. A board game has to have clear rules, provide a sufficient rationale for collaboration, be a challenge for participants, and provide different types of activities and experiences. It is connected with the next important step, which is to precisely define the goals of the event, regarding prepared tasks. Chosen activities should allow assessment not only the content knowledge but also all other aspects (e.g., Science as a Human Endeavor and Science Inquiry Skills) chosen by the teacher. Breedlove et al. (2004) reported that the effects of collaborative testing were directly related to the level of cognitive processing required by the test questions. The activities, rules, and scoring system should be modified and matched to the groups. Particularly, in our research, one student claimed that there was a possibility to guess the proper word in the charades without physics knowledge. This can be improved by additional rules or modifying the charades questions with other types of activities. Because the effectiveness of the collaborative testing may depend on students' earlier teaching strategies and improve over time as students become more familiar with the collaborative process (Castor 2004), the modification of the game seems to be a natural consequence.

Further Issues

The study examined only the short-term effect on students' retention knowledge. One can suppose that because the initial level of the forgetting curve was higher in experimental groups than in control groups, after a few months, experimental groups should also obtain better results. This assumption has to be, however, tested in future work. The method could be also implemented and verified in subjects other than physics as well as in a wider spectrum of school levels. Even though literature findings about collaborative testing and board games in many science subjects are very enthusiastic, only a few of them focus on the distinction between high- and low-performing students (Giuliodori et al. 2008). This question should be also examined in future work.

Another issue is the claim that a teacher could influence the results, e.g., by focusing on the experimental group or neglecting control groups. In our research, control group results in the post-test were similar to their results in all other tests taken before, taken in average as former achievements. This approach, unlike the typical pre-test, allows us to address this remark. Comparison between pre- and post-test provides clear information about students' gain in the examination topic, but can be easily influenced by the teacher, which will be found only in the lower achievements for the control group. In our approach, we assumed that an uninfluenced teaching style will have an effect in unchanged students' results in the post-test, which is assumed to be allowed. However, the implementation of the method under different circumstances could also provide worthwhile information.

Future research could also examine collaborative testing as a more effective standard assessment strategy across a curriculum (Meseke 2010). Because the game always has to be a challenge for students, some modifications should be introduced in the type of questions and rules or the board game should be used interchangeably with another collaborative method.

Concluding Remarks

This paper studied the influence of a board game as an assessment method on high school students' achievement. Students from experimental groups performed better in the game than in the former tests. Simultaneously their achievement in a traditional test taken 1 week after was significantly higher than for students from control groups. It implies that assessing students' achievement in the form of a game may improve their performance and short-term achievements.

The improvement of students' achievements may result in combining collaborative testing with gamification. Apart from quantitative results, the students' enthusiastic opinions are also indicative of the social benefits of the approach, such as the development of group work skills, supporting weaker students through collaboration with others, and, in addition to these, integration of the class. It appears that game-based assessment enhances students' retention of knowledge and provides opportunities for improvement for each student, regardless of their former performance. Moreover, it helps to improve students' attitudes towards their learning and add valuable collaborative learning experience to enhance the school curriculum.

The approach can be easily modified and adapted as a testing method in fields other than physics, especially natural sciences, in which assessing the experimental skills and socio-historical context are also under consideration.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Adams, P. (2006). Exploring social constructivism: Theories and practicalities. *Education*, 34(3), 243–257.
- Akl, E. A., Kairouz, V. F., Sackett, K. M., Erdley, W. S., Mustafa, R. A., Fiander, M., Gabriel, C., & Schunemann, H. (2013). Educational games for health professionals. *Corhane Database of Systematic Reviews*, 3, 1–46.
- Allery, L. (2004). Educational games and structured experiences. *Medical Teacher*, 26(6), 504–505.
- Apostol, S., Zaharescu, L., & Aleze, I. (2013). Gamification of learning and educational games. *eLearning & Software for Education*, 2, 67.
- Banfield, J., & Wilkerson, B. (2014). Increasing student intrinsic motivation and self-efficacy through gamification pedagogy. *Contemporary Issues in Education Research CIER*, 7(4), 291.
- Bloom, D. (2009). Collaborative test taking. *College Teaching*, 57(4), 216–220.
- Bovee, M., & Gran, D. (2005). Effects of collaborative testing on students satisfaction survey. *Journal of Chiropractic Education*, 19(1), 47.
- Bragg, L. A. (2003). Children's perspectives on mathematics and game playing. In L. Bragg, C. Campbell, G. Herbert, & J. Mousley (Eds.), *Mathematics Education Research: Innovation, Networking, Opportunity: Proceedings of the 26th Annual Conference of the Mathematics Education Research Group of Australasia* (Vol. Vol. 1, pp. 160–167). Geelong: Deakin University.
- Bragg, L. A. (2006). *The impact of mathematical games on learning, attitudes, and Behaviours*. Bundoora: La Trobe University.
- Breedlove, W., Burkett, T., & Winfield, I. (2004). Collaborative testing and test performance. *Academic Exchange Quarterly*, 8(3), 36–40.
- Bright, G. W., Harvey, J. G., & Wheeler, M. M. (1983). Use of a game to instruct on logical reasoning. *School Science and Mathematics*, 83(5), 396–405.
- Burguillo, J. C. (2010). Using game theory and competition-based learning to stimulate student motivation and performance. *Computers & Education*, 55(2), 566–575.

- Castor, T. (2004). Making student thinking visible by examining discussion during group testing. *New Directions for Teaching and Learning*, 100, 95–99.
- Conklin, A. (2006). Cyber defense competitions and information security education: An active learning solution for a capstone course. In System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on (Vol. 9, pp. 220b–220b). IEEE.
- Cortright, R. N., Collins, H. L., Rodenbaugh, D. W., & Dicarolo, S. E. (2003). Student retention of course content is improved by collaborative-group testing. *Advances in Physiology Education*, 27(3), 102–108.
- Dallmer, D. (2004). Collaborative test taking with adult learners. *Adult Learning*, 15, 4–7.
- Damon, W., & Phelps, E. (1989). Critical distinctions among three methods of peer education. *International Journal of Educational Research*, 13, 9–19.
- Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification: Using game-design elements in non-gaming contexts. In Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems (p. 2425). Vancouver, Canada.
- Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Educational Technology & Society*, 18(3), 75–88.
- Dienes, Z. P. (1963). *An experimental study of mathematics learning*. New York: Hutchinson.
- Doolittle, P. (1997). Vygotsky's zone of proximal development as a theoretical foundation for cooperative learning. *Journal on Excellence in College Teaching*, 8(1), 83–103.
- Duane, B. T., & Satre, M. E. (2014). Utilizing constructivism learning theory in collaborative testing as a creative strategy to promote essential nursing skills. *Nurse Education Today*, 34(1), 31–34.
- Durden, T., & Dangel, J. (2008). Teacher-involved conversations with young children during small group activity. *Early Years*, 28, 251–266.
- Dziob, D., Kwiatkowski, L., & Sokolowska, D. (2018). Introducing a class tournament as an assessment method of student achievements in physics courses. *EURASIA Journal of Mathematics, Science and Technology Education*, 14(4), 1111–1132.
- Ernest, P. (1986). Games: A rationale for their use in the teaching of mathematics in school. *Mathematics in School*, 15(1), 2–5.
- Eskelinen, M. (2001). Towards computer game studies. *Digital Creativity*, 12(3), 175–183.
- Gilley, B., & Clarkston, B. (2014). Research and teaching: Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students. *Journal of College Science Teaching*, 043(03), 83.
- Giuliodori, M. J., Lujan, H. L., & DiCarlo, S. E. (2008). Collaborative group testing benefits high- and low-performing students. *Advances in Physiology Education*, 32(4), 274–278.
- Griffin, S. (2004). Building number sense with number worlds: A mathematics program for young children. *Early Childhood Research Quarterly*, 19, 173–180.
- Guest, K. E., & Murphy, D. S. (2000). In support of memory retention: A cooperative oral final exam. *Education*, 121, 350–354.
- Haberyan, A., & Barnett, J. (2010). Collaborative testing and achievement: Are two heads really better than one? *Journal of Instructional Psychology*, 37(1), 32–41.
- Hanus, M., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, 80, 152–161.
- Harlen, W., & James, M. (1997). Assessment and learning: Differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365–379.
- Hendrix, J. C. (1996). Cooperative learning: Building a democratic community. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 69(6), 333–336.
- Hodges, L. (2004). Group exams in science courses. *New Dir Teach Learn*, 2004, 89–93.
- Ifenthaler, D., Eseryel, D., & Ge, X. (2012). *Assessment in game-based learning: Foundations, innovations, and perspectives*. New York: Springer.
- Jensen, M., Moore, R., & Hatch, J. (2002). Cooperative learning: Part I: Cooperative quizzes. *The American Biology Teacher*, 64(1), 29–34.
- Jolliffe, W. (2007). *Cooperative learning in the classroom: Putting it into practice*. London: Paul Chapman.
- Kapitanoff, S. H. (2009). Collaborative testing: Cognitive and interpersonal processes related to enhanced test performance. *Active Learning in Higher Education*, 10(1), 56–70.
- Ko, S. (2002). An empirical analysis of children's thinking and learning in a computer game context. *Educational Psychology*, 22(2), 219–233.
- Leight, H., Saunders, C., Calkins, R., & Withers, M. (2012). Collaborative testing improves performance but not content retention in a large-enrollment introductory biology class. *CBE Life Sciences Education*, 11(4), 392–401.
- Ligeikis-Clayton, C. (1996). Shared test taking. *J NY State Nurses Assoc*, 27(4), 4–6.
- Looney, J. (2011). Integrating Formative and Summative Assessment. OECD Education Working Papers, 58(5).

- Lusk, M., & Conklin, L. (2003). Collaborative testing to promote learning. *Journal of Nursing Education*, 42(3), 121–124.
- McTighe, J., & O'Connor, K. (2005). Seven practices for effective learning. *Educational Leadership*, 63(10), 10–17.
- Meseke, C., Nafziger, R., & Meseke, J. (2008). Student course performance and collaborative testing: A prospective follow-on study. *Journal of Manipulative and Physiological Therapeutics*, 31(8), 611–615.
- Meseke, C. A., Bovee, M. L., & Gran, D. F. (2009). The impact of collaborative testing on student performance and satisfaction in a chiropractic science course. *Journal of Manipulative and Physiological Therapeutics*, 32, 309–314.
- Meseke, C., Nafziger, R., & Meseke, J. (2010). Student attitudes, satisfaction, and learning in a collaborative testing environment*. *Journal of Manipulative and Physiological Therapeutics*, 24(1), 19–29.
- Michaelson, L. K., Knight, A. B., & Fink, D. L. (2002). *Team-based learning: A transformative use of small groups*. Westport, CT: Praeger.
- Millis, B. J., & Cottell, P. G. (1998). *Cooperative learning for higher education faculty*. Phoenix: American Council on Education and Oryx Press.
- Moncada, S., & Moncada, T. (2014). Gamification of learning in accounting education. *Journal of Higher Education Theory and Practice*, 14, 9.
- Neame, R., & Powis, D. (1981). Toward independent learning: Curricular design for assisting students to learn how to learn. *Journal of Medical Education*, 56(11), 886–893.
- Oldfield, B. J. (1991). Games in the learning of mathematics—Part 1: Classification. *Mathematics in School*, 20(1), 41–43.
- Ramani, G. B., & Siegler, R. S. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science*, 11(5), 655–661.
- Richardson, D., & Birge, B. (1995). Teaching physiology by combined passive (pedagogical) and active (andragogical) methods. *The American Journal of Physiology*, 268, S66–S74.
- Rieber, L., & Noah, D. (2008). Games, simulations, and visual metaphors in education: Antagonism between enjoyment and learning. *Educational Media International*, 45(2), 77–92.
- Royse, M. A., & Newton, S. E. (2007). How gaming is used as an innovative strategy for nurse education. *Nursing Education Perspectives*, 28(5), 263–267.
- Sadler, T. D., Romine, W. L., Stuart, P. E., & Merle-Johnson, D. (2013). Game-based curricula in biology classes: Differential effects among varying academic levels. *Journal of Research in Science Teaching*, 50(4), 479–499.
- Sandahl, S. S. (2010). Collaborative testing as a learning strategy in nursing education. *Nursing Education Perspectives*, 31(3), 142–147.
- Saunders, N. A., & Wallis, B. J. (1981). Learning decision-making in clinical medicine: A card game dealing with acute emergencies for undergraduate use. *Medical Education*, 15(5), 323–327.
- Sawtelle, V., Brewé, E., & Kramer, L. H. (2012). Exploring the relationship between self-efficacy and retention in introductory physics. *Journal of Research in Science Teaching*, 49(9), 1096–1121.
- Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, 74, 14–31.
- Shanklin, S. B., & Ehlen, C. R. (2007). Using the monopoly board game as an in-class economic simulation in the introductory financial accounting course. *Journal of College Teaching & Learning*, 4(11), 17–22.
- Shindler, J. V. (2003). "Greater than the sum of the parts?" examining the soundness of collaborative exams in teacher education courses. *Innovative Higher Education*, 28(4), 273–283.
- Simpkin, M. G. (2005). An experimental study of the effectiveness of collaborative testing in an entry-level computer programming class. *Journal of Information Systems*, 16, 273–280.
- Slusser, S. R., & Erickson, R. (2006). Group quizzes: An extension of the collaborative learning process. *Teaching Sociology*, 34(3), 249–262.
- Steinman, R. A., & Blastos, M. T. (2002). A trading-card game teaching about host defence. *Medical Education*, 36(12), 1201–1208.
- Sullivan, P. (1993). Short flexible mathematics games. In J. Mousley & M. Rice (Eds.), *Mathematics of primary importance* (pp. 211–217). Brunswick, Victoria: The Mathematical Association of Victoria.
- Sung, H., & Hwang, G. (2013). A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers & Education*, 63, 43–51.
- Talanquer, V., Bolger, M., & Tomanek, D. (2015). Exploring prospective teachers' assessment practices: Noticing and interpreting student understanding in the assessment of written work. *Journal of Research in Science Teaching*, 52(5), 585–609.
- Wasik, B. A. (2008). When fewer more: Small groups in early childhood classrooms. *Early Childhood Education Journal*, 35, 515–521.

- Wiliam, D., & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22(5), 537–548.
- Wininger, S. (2005). Using your tests to teach: Formative summative assessment. *Teaching of Psychology*, 32(3), 164–166.
- Woody, W., Woody, L. & Bromley, S. (2008). Anticipated group versus individual examinations: a classroom comparison. *Teach Psychol.* 35, 13–17.
- Zagal, J., Rick, J., & Hsi, I. (2006). Collaborative games: Lessons learned from board games. *Simulation & Gaming*, 37(1), 24–40.
- Zimbardo, P. G., Butler, L. D., & Wolfe, V. A. (2003). Cooperative college examinations: More gain, less pain when students share information and grades. *The Journal of Experimental Education*, 71(2), 101–125.